Andrew Chu

647-966-2880 | andrew.chu206@gmail.com | /in/andrewchu-ca | github.com/andrewchu16 | andrewchu.ca

Education

University of Waterloo May 2029
Bachelor of Computer Science (Honours) *Waterloo, ON*

Experience

Software Engineer Intern May 2025 - Aug. 2025
CGI *Markham, ON*

- Developed multimodal RAG system to speed up information retrieval from **500+** internal production guides, accelerating team incident response times by **17%**
- Automated health checks, server restarts, and email notification services, by engineering unified dashboard for remotely executing jobs on **400+** applications and servers with Ansible
- Increased chat response accuracy by **38%** through PEFT fine-tuning a Llama3 model with QLoRA

Technical Lead (Part Time) May. 2025 - Aug. 2025
University of Waterloo's Data Science Club *Waterloo, ON*

- Spearheaded user check-in and application portal projects written in Next.js and Express with **1000+** users
- Resolved **15+** development tickets spanning frontend UI enhancements, application submission workflows, and backend database migration
- Implemented email confirmation for password resets and user portal, improving user experience and security

Founding Engineer May. 2025 - Present
Villio AI *Waterloo, ON*

- Architectured FastAPI backend for multi-tenant city services app, integrating Firebase Auth for user authentication
- Engineered agentic pipeline using LangChain for filling out forms through a chat interface for ease of use
- Deployed backend securely through Azure App Service, enabling high availability across multiple cities

Data Engineer Intern Aug. 2024 - Sep. 2024
Alljoined *Toronto, ON*

- Debugged critical timing issue in data preprocessing by implementing data observability system in Python
- Improved data collection script performance by **20x** by refactoring WebSocket script to prevent deadlocking in network communications between EEG headset and compute cluster
- Designed data processing pipeline for denoising and artifact correction using NumPy, Pandas, and scikit-learn

Full-Stack Developer Sep. 2022 - Jan. 2023
Digitera Interactive *Ottawa, ON*

- Developed "Skule News" project, a mobile school news app to streamline student communication written in Flutter
- Integrated Firebase for authentication, real-time updates, and cloud file storage for secure backend support
- Reduced frequency of in-app crashes by **20%** by resolving **6+** bugs resulting from incorrect Firebase queries

Projects

Replate - Sustainable Delivery App (GenAI Genesis Winner) | *Next, Express, MongoDB, Flask, Cohere, Gemini, Twilio*

- Achieved **AI Eco-Mobility Award** at GenAI Genesis, Canada's largest AI hackathon, among **600+** competitors
- Engineered a meal recommendation system using an evaluator-optimizer agentic workflow with Cohere and Gemini, combining user preferences with restaurant data from the *Too Good to Go* platform
- Utilized Express and Flask backend microservices to process orders and serve live delivery updates via Twilio SMS

chat.andrewchu.ca - Personal AI Assistant | *Transformers, Nginx, PyTorch, mlx-lm, Docker, PostgreSQL, Next, FastAPI*

- Built a personal AI chatbot fine-tuned using mlx-lm library using QLoRA, deployed on a Raspberry Pi
- Cut LLM inference times by **67%** by embedding queries with mxbai-embed-xsmall-v1 and caching LLM answers with HNSW vector similarity search

Technical Skills

Languages: TypeScript, Python, C/C++, Java, Go, C#, CSS, HTML, Bash
Frameworks: Next, React, Express, Flask, Redis, Jest, SvelteKit, Electron, LangChain, MySQL, MongoDB, Supabase
Developer Tools: Azure, WebLogic, Linux, Git, Figma, Docker, Postman, GCP, Nginx